working paper
department
of economics

The Inconsistency of Distributed Lag Estimators
Due to Misspecification by Time Aggregation

Robert F. Engle

Massachusetts Institute of Technology

Number 63 October 1970

massachusetts
institute of
technology

50 memorial drive
cambridge, mass. 02139

The Inconsistency of Distributed Lag Estimators Due to Misspecification by Time Aggregation

Robert F. Engle

Massachusetts Institute of Technology

Number 63 October 1970

As econometricians find themselves capable of obtaining and handling data which is disaggregated over time, it becomes necessary to compare the specification of models which are formulated and estimated in different unit time periods. For static models, time aggregation is not a problem as the model can be aggregated by successively lagging and then summing the equations. Nerlove [14], Ironmonger [9], and Mundlank [13] early explored the problem in dynamic models and showed that it is a complicated process. Sims [16] has analyzed the best fit of a discrete lag distribution to a continuous time underlying model but with little regard for the estimation problems. Telser [17] on the other hand derived an estimator for the true underlying coefficients in terms of the aggregated data by extracting the missing information from the process of the aggregated residual. His procedure is however only asymptotically unique and not very robust, and consequently is not a practical estimation procedure. This paper will show, for a series of underlying model specifications, what the asymptotic values of ordinary least squares or two stage least squares estimates of the parameters in the aggregated model will be. The purpose is therefore to enable an experimenter to compare different levels of aggregation or to anticipate the biases from failing to use sufficiently disaggregated data. To check the validity of the approximations made in the analysis, seven of Liu's [11]monthly model equations are aggregated to quarterly and annual forms and reestimated.

I. Methodology

Two distributed lag models which are formulated in different unit periods can be compared only if unit invarient statistics of the two distributions can be found. Two such statistics recommended by Griliches [6]

are the long run propensity and the average lag. Never can the lag distributions be exactly the same, but if these two simple parameters are the same, we might be sanguine about our estimates. The procedure will therefore be to employ Theil's theory of specification analysis to discover what variations in these two parameters might be expected to result from time aggregation of a distributed lag model.

Defining the disaggregated time unit as a month and using L as the lag operator, the underlying true model is in general given by

$$1) \qquad \alpha(L)y = \beta(L)x + \gamma(L)\varepsilon$$

where ε is independent of x; $\alpha(L)$, $\beta(L)$ and $\gamma(L)$ are all rational polynomials in L; and x and y are measured as deviations from sample means. Aggregated data, which will always be capitalized, is obtained from the monthly data by applying an aggregation operator $R(L)$ as follows

$$2) \qquad Y_t = R(L)y_t \qquad\qquad t = n, 2n, 3n, \ldots$$

$$X_t = R(L)x_t,$$

where $R(L) = \frac{1}{n}(1 + L + L^2 + \ldots + L^{n-1})$ if x and y are flow variables expressed in annual rates, or where $R(L) = 1$ if x and y are stock variables defined at the end of the period. Notice that for standard annual data Y_t, X_t are only defined for $t = 12, 24, 36, \ldots$ since t counts in months. Clearly the stock variable case is easier but it is in general not as important for econometric applications. The aggregate model which has been most carefully analyzed in this paper is the Koyck-Nerlove model

$$3) \qquad Y_t = AY_{t-n} + BX_t + W_t \qquad t = 2n, 3n, \ldots$$

which is almost the simplest lag model and is certainly one of the most important. In principle, however, any model could have been used but the computations would have been more involved.

By writing $\text{plim} \frac{1}{T} x'y \equiv (x,y)$ we can conveniently write the probability limit of the estimated parameters as

$$4) \quad \text{plim} \begin{pmatrix} \hat{A} \\ \hat{B} \end{pmatrix} = \begin{pmatrix} (Y_{-n},Y_{-n}) & (Y_{-n},X) \\ (Y_{-n},X) & (X,X) \end{pmatrix}^{-1} \begin{pmatrix} (Y_{-n},Y) \\ (X,Y) \end{pmatrix}$$

where t has been suppressed. Assuming that the processes are covariance stationary this becomes

$$5) \quad \text{plim} \begin{pmatrix} \hat{A} \\ \hat{B} \end{pmatrix} = \begin{pmatrix} \dfrac{(X,X)\ (Y,Y_{-n}) - (Y_{-n},X)\ (Y,X)}{(Y,Y)\ (X,X) - (Y_{-n},X)^2} \\ \\ \dfrac{(Y,Y)\ (Y,X) - (Y_{-n},X)\,(Y_{-n},Y)}{(Y,Y)\ (X,X) - (Y_{-n},X)^2} \end{pmatrix}$$

There are only five moments in (5); once these are computed the probability limits of the aggregated coefficients are known and the Av. Lag. $= \frac{nA}{1-A}$ and the long run marginal propensity $\frac{B}{1-A}$ are easily calculated. The difficulty is in computing these moments under the assumption that (1) and (2) provide the true stochastic specification of the variables.

Much of the analysis can in principle be performed either with the conventional methods of difference equations or in the spectral representation. As the latter is far more versatile and computationally simpler it will be used exclusively in this paper. For a simple example of the difference equation methodology see Engle and Liu [3] or Engle [2] .

If we let the power spectrum of x be $f_x(\theta)$ and the power spectrum of ε be flat with height σ^2 which means that ε is white noise, then using spectral transfer functions*, we can express the desired moments as integrals.**

*For examples of this see Nerlove [15] or Howrey [8] .

**In addition we must assume that the moments are consistent estimates of covariances which will be true if we assume that higher order moments are finite, Goldberger [4] , or that the process is Gaussion, Doob [1] , or in general that the process exhibits asymptotic independence.

For example, the variance of x is

$$6)\qquad (x,x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_x(\theta)\, d\theta$$

and so

$$7)\qquad (X,X) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|R\ (e^{i\theta})\right|^2 f_x(\theta)\, d\theta$$

since

$$8)\qquad f_X(\theta) = \left|R(e^{i\theta})\right|^2 f_x(\theta)\ ^*$$

where the role of $\left|R(e^{i\theta})\right|^2$ in (7) is that of a spectral transfer function. Using the same arguments we can write the following series of relations using (1) and (2)

$$f_y(\theta) = \left|\frac{\beta(e^{i\theta})}{\alpha(e^{i\theta})}\right|^2 f_x(\theta) + \left|\frac{\gamma(e^{i\theta})}{\alpha(e^{i\theta})}\right|^2 \sigma^2$$

$$f_{yx}(\theta) = \frac{\beta(e^{-i\theta})}{\alpha(e^{-i\theta})} f_x(\theta)$$

$$9)\qquad f_Y(\theta) = \left|R(e^{i\theta})\right|^2 \left|\frac{\beta(e^{i\theta})}{\alpha(e^{i\theta})}\right|^2 f_x(\theta) + \left|R(e^{i\theta})\right|^2 \left|\frac{\gamma(e^{i\theta})}{\alpha(e^{i\theta})}\right|^2 \sigma^2$$

$$f_{YX}(\theta) = \left|R(e^{i\theta})\right|^2 \frac{\beta(e^{-i\theta})}{\alpha(e^{-i\theta})} f_x(\theta)$$

---

*$f_X(\theta)$ is actually the power spectrum of X only if it is observed monthly. If instead X can only be observed for example annually, then the higher frequency components cannot be separated from some of the low frequencies. In fact, the power spectrum is only defined for $-\pi/12\leq\theta\leq\pi/12$ and all θ's outside this interval are "folded" into this interval. However, the integral in (7) should in fact only cover the interval $-\pi/12\leq\theta\leq\pi/12$ and the result is identical. This result is simply that a consistent estimate of the annual variance is obtained either by using every monthly observation of the annual average of just every twelfth one.

and from (9) we can write the expressions for the required moments directly.

$$(Y,Y) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|R(e^{i\theta})\right|^2 \left|\frac{\beta(e^{i\theta})}{\alpha(e^{i\theta})}\right|^2 f_x(\theta)\, d\theta$$

$$+ \frac{1}{2\pi}\int_{-\pi}^{\pi} \sigma^2 \left|R(e^{i\theta})\right|^2 \left|\frac{\gamma(e^{i\theta})}{\alpha(e^{i\theta})}\right|^2 d\theta$$

$$(Y_{-n},Y) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-in\theta} \left|R(e^{i\theta})\right|^2 \left|\frac{\beta(e^{i\theta})}{\alpha(e^{i\theta})}\right|^2 f_x(\theta)\, d\theta$$

10)

$$+ \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-in\theta} \sigma^2 \left|R(e^{i\theta})\right|^2 \left|\frac{\gamma(e^{i\theta})}{\alpha(e^{i\theta})}\right|^2 d\theta$$

$$(Y,X) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|R(e^{i\theta})\right|^2 \frac{\beta(e^{-i\theta})}{\alpha(e^{-i\theta})} f_x(\theta)\, d\theta$$

$$(Y_{-n},X) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-in\theta} \left|R(e^{i\theta})\right|^2 \frac{\beta(e^{-i\theta})}{\alpha(e^{-i\theta})} f_x(\theta)\, d\theta$$

Thus by choosing $\alpha(L)$, $\beta(L)$, $\gamma(L)$, $f_x(\theta)$ and σ^2 we can obtain the probability limit of the coefficients of any level of aggregation just by evaluating these integrals.

Unfortunately, as the reader might suspect, this is not a trivial step. In practice, the simplest method for performing this integration is to use Cauchy's Integral Theorem which is explained and employed in the appendix. With this tool the biases are now computable.

II. Analytic Results

The specific underlying model examined is

11) $$(1-\alpha L)\quad y = \beta x + \frac{(1-\delta L)}{(1-\gamma L)}\,\varepsilon$$

where ε is white noise with variance σ^2. Special cases of this model

cover most of the controversies over the estimation procedure for Koyck-Nerlove adjustment models; that is $\delta = \gamma$ is the serially uncorrelated disturbance which allows direct estimation of the model; $\delta = 0$ suggests a simple autoregressive transformation for estimation; $\delta = \alpha$, $\gamma = 0$ is the moving average disturbance which occurs when the uncorrelated disturbance is introduced before the Koyck transformation; and $\delta = \alpha$, $\gamma \neq 0$ is the example where both moving average and autoregressive processes exist in the disturbance.

Within the structure of this model, a power spectrum for x must be assumed. Two candidates have been considered. One is that x is a first order autoregressive process with parameter ρ; while the other is that x is a purely deterministic oscillating process with frequency ω. The former assumption has been analyzed in Engle [2] and proves to give more complicated results than the latter; but under the approximation that ρ is close to unity (that the variable does not vary much from one month to another) the conclusions are identical with the $\omega = 0$ approximation discussed below.

The assumption of a deterministic process with frequency ω turns out to be a fruitful approach for economic time series. It is equivalent to assuming that the spectrum is zero everywhere but has a spike at $\pm\,\omega$ with height σ_x^2. Clearly, this makes many of the integrals simpler and in particular, the case where $\omega = 0$ is a reasonable approximation to many economic time series which have important trends. Granger [5] argues that typical economic variables have most of their power in very low frequencies and Hannan and Terrell [7] use the same approximation. Furthermore, since many time series have a series of peaks corresponding to seasonal harmonics, a superposition of these deterministic

components each weighted by a reasonable variance, would give a good representation of a time series with strong seasonality.

The probability limit of the estimate of the parameter A in the aggregated model (3) when x is deterministic is given by the following unpleasant expression:

$$\text{(12)}\quad \text{plim}\,\hat{A} = \frac{\dfrac{MG(\omega)}{T(\omega)}\left[\dfrac{N(\alpha)\;\alpha(1-\alpha\delta)\;(\alpha-\delta)\;(1-\gamma^2) - N(\gamma)\;\gamma(1-\gamma\delta)\;(\gamma-\delta)\;(1-\alpha^2)}{(1-\alpha^2)\;(1-\gamma^2)\;(\alpha-\gamma)\;(1-\alpha\gamma)}\right] + \cos n\omega - \dfrac{(1-\alpha\cos\omega)\;(\cos n\omega - \alpha\cos(n-1)\omega)}{G(\omega)}}{\dfrac{MG(\omega)}{T(\omega)}\left[\dfrac{Q(\alpha)\;(1-\alpha\delta)\;(\alpha-\delta)\;(1-\gamma^2) - Q(\gamma)\;(1-\gamma\delta)\;(\gamma-\delta)\;(1-\alpha^2)}{(1-\alpha^2)\;(1-\gamma^2)\;(\alpha-\gamma)\;(1-\alpha\gamma)}\right] + 1 - \left(\dfrac{\cos n\omega - \alpha\cos(n-1)\omega}{G(\omega)}\right)^2}$$

where

$$N(\alpha) = \frac{1}{n^2}\left[1 + 2\alpha + 3\alpha^2 + \ldots + n\alpha^{n-1} + (n-1)\alpha^n + \ldots + \alpha^{2(n-1)}\right]$$

$$Q(\alpha) = \frac{1}{n^2}\left[n + 2(n-1)\alpha + 2(n-2)\alpha^2 + \ldots + 2\alpha^{n-1}\right]$$

$$G(\omega) = 1 + \alpha^2 - 2\alpha\cos\omega$$

$$T(\omega) = \frac{1}{n^2}\left[n + 2(n-1)\cos\omega + 2(n-2)\cos 2\omega + \ldots + 2\cos(n-1)\omega\right]$$

$$M = \sigma_\varepsilon^2 / \beta^2\sigma_x^2$$

Note that for $\alpha \geq 0$, $N(\alpha) \leq Q(\alpha) \leq 1$ with equality for $\alpha = 1$, $(1-\alpha)^2 \leq G(\omega) \leq (1+\alpha)^2$, and $-1 \leq T(\omega) \leq 1$. The terms in square brackets come directly from the process of the disturbance while the other four terms result from the process of the exogenous variable.

There are two important cases in which we will find the disturbance terms most important. First, if M (which is $1-R^2/R^2$ of the monthly model) is large,

then the other terms which are all less than one will be unimportant. Secondly if the frequency of the exogenous variable is approximately zero, then only the disturbance terms will be left. Since all terms are cosines, the derivative with respect to ω is zero at $\omega = 0$ and so there is no abrupt change in the validity of the approximation if ω is merely close to zero. In addition, suppose there were a second peak in the spectrum at a frequency ω, different from zero with a Variance σ_1^2 somewhat less than σ_0^2 which corresponds to $\omega = 0$. Then there would be three more terms in the numerator and denominator of (12) where each of the new disturbance terms would be weighted by $\dfrac{\sigma_1^2\, T(\omega_1)}{\sigma_0^2\, T(0)}$, and the other new terms by $\dfrac{\sigma_1^4\, T^2(\omega_1)}{\sigma_0^4\, T^2(0)}$.

Since $T(\omega)$ is a maximum of 1 and this occurs when $\omega = 0$, the new terms would find their contribution receiving little weight. However, it is of course true that the terms themselves might be very large and thus overcome the effect of small weights. Rarely does a power spectrum exhibit a peak higher than that at $\omega = 0$, especially if it has not been carefully detrended.

This argument suggests that a careful examination of (12) when $\omega = 0$ would often give approximation to much of economic reality. The result is

$$13)\qquad \text{plim}\ \hat{A} = \frac{N(\alpha)\alpha(1-\alpha\delta)(\alpha-\delta)(1-\gamma^2) - N(\gamma)\gamma(1-\gamma\delta)(\gamma-\delta)(1-\alpha^2)}{Q(\alpha)\ (1-\alpha\delta)(\alpha-\delta)(1-\gamma^2) - Q(\gamma)\ (1-\gamma\delta)(\gamma-\delta)(1-\alpha^2)}.$$

The simplest model is of course obtained when $\delta = \gamma = 0$, implying that the true model is

$$14)\qquad y = \alpha y_{-1} + \beta x + \varepsilon$$

where the disturbance has a white noise process. In this case (13) becomes:

$$15)\qquad \text{plim}\ A = \frac{\alpha N(\alpha)}{Q(\alpha)}$$

which is always less than or equal to x and greater than or equal to α^n. However, $\frac{n \text{ plim } \hat{A}}{1-\text{plim } A} - \frac{\alpha}{1-\alpha}$ does not have a unique sign. Generally, it rises for $n < 4$ and then falls becoming negative. That is, while a quarterly model might overestimate the average lag, an annual model would underestimate it.

When $\delta = \alpha$ and $\gamma = 0$, the true monthly model is:

$$16) \qquad y = \alpha y_{-1} + \beta x + \varepsilon - \alpha\varepsilon_{-1}$$

with a moving average disturbance which results from an atuoregressive transformation of a model with white noise disturbances. The expression for the estimator is

$$17) \qquad \text{plim } \hat{A} = \frac{\gamma N(\gamma)}{Q(\gamma)} .$$

Thus the observed lag structure has nothing to do with the true x, only with the original serial correlation. The estimated average lag will however rise and then fall with aggregation just as in (14). The reason for the starkness of this result, is that our assumption about the process of the exogenous variable emasculates the lag distribution when it is expressed as an infinite moving average with geometrically declining weights. Nevertheless, (17) provides an important contribution to the estimate of A and the direction and origin of this bias should be clear.

Perhaps the most important case is when the disturbance has a first order markov process and the original model is

$$18) \qquad y = \alpha y_{-1} + \beta x + \frac{\varepsilon}{1-\gamma L} .$$

This model is in a form suitable for various estimation techniques and the question of estimation with serially correlated disturbances is generally posed within the context of (18). The moving average disturbance is more difficult to handle and usually is just approximated by a rapidly decaying autoregressive process. The estimator of A resulting from (18) is

$$19)\qquad \text{plim}\ \hat{A} = \frac{N(\alpha)\ \alpha^2(1-\gamma^2) - N(\gamma)\ \gamma^2(1-\alpha^2)}{Q(\alpha)\ \alpha\ (1-\gamma^2) - Q(\gamma)\ \gamma\ (1-\alpha^2)}\ .$$

To see the biases in the average lag as estimated with (19) a plot of the average lag for $\gamma = .5, 0, -.5$ is plotted against the number of periods of aggregation for $\alpha = .55$ and .9 in figures 1 and 2. In addition, these curves are replotted for $\omega = .05$ and .1 with $M = .5$ in each case to examine the effect on the estimates of small deviations from $\omega = 0$. These ω's correspond to oscillations of 125 and 63 periods respectively.

Several important facts can be seen by studying these two figures. First, the effect of serial correlation is strongest in relatively disaggregated models and leads to an overall increase or decrease in the estimated average lag depending on whether the serial correlation is negative or positive respectively. Second, in the more highly aggregated models, the size of the serial correlation coefficient seems less important than the deviation of ω from zero in explaining the size of the estimate of the average lag. In particular the larger is ω the larger will be the aggregated average lag. If ω gets even larger, other calculations show that the average lag goes to infinity and then becomes negative and large. This results from the fact that the denominator of (12) goes through zero as ω increases, but it seems unlikely that this effect will be of practical importance.

The third important observation is that the consequences on the estimated average lag of having $\omega \neq 0$ seem to be much stronger when α is close to one. In a monthly model one would expect the true α to be very large and thus deviations from $\omega = 0$ to have a strong effect. It must be remembered, however, that all these calculations were made under the assumption that $M = .5$ and therefore any comparision between

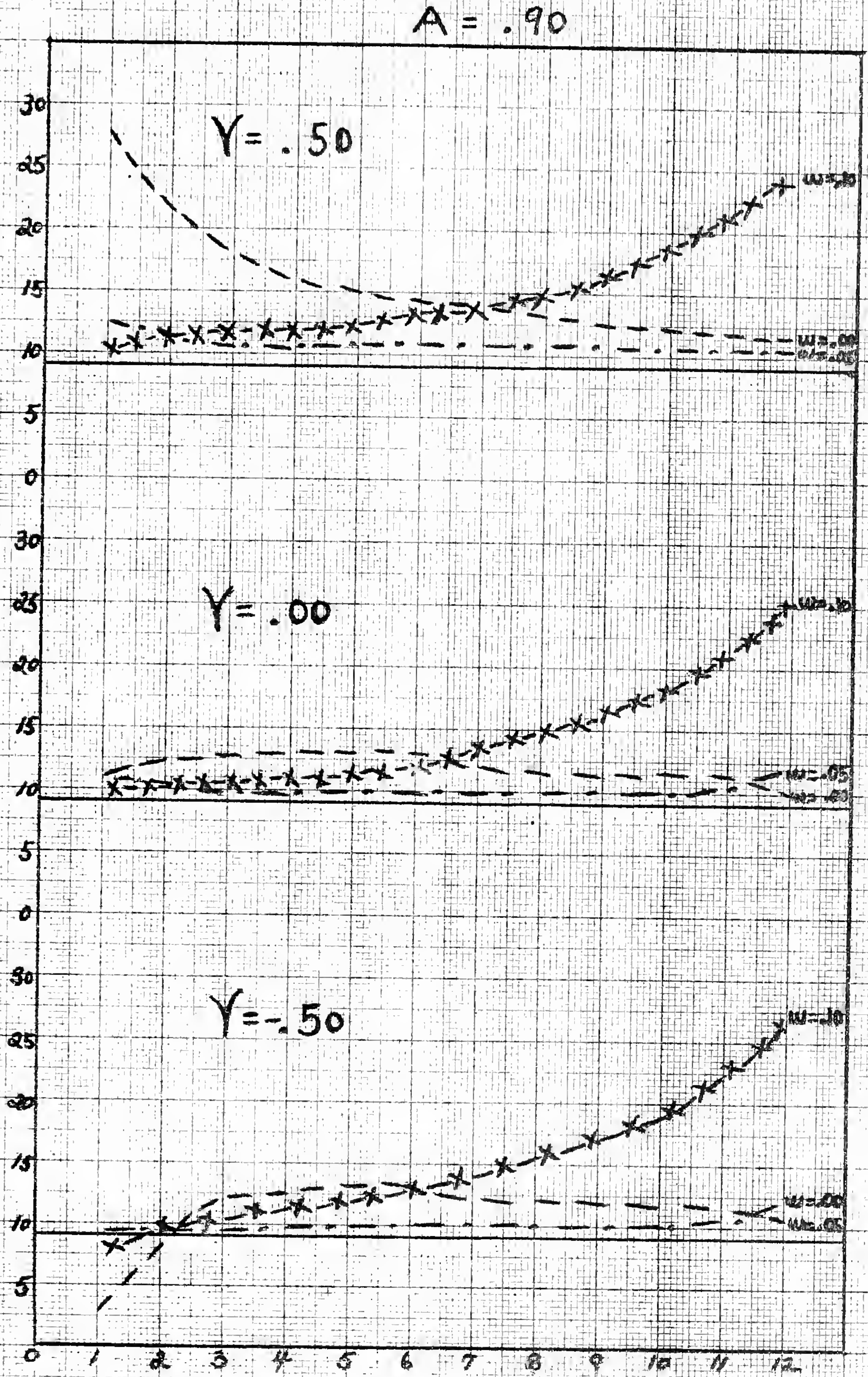

figure 2

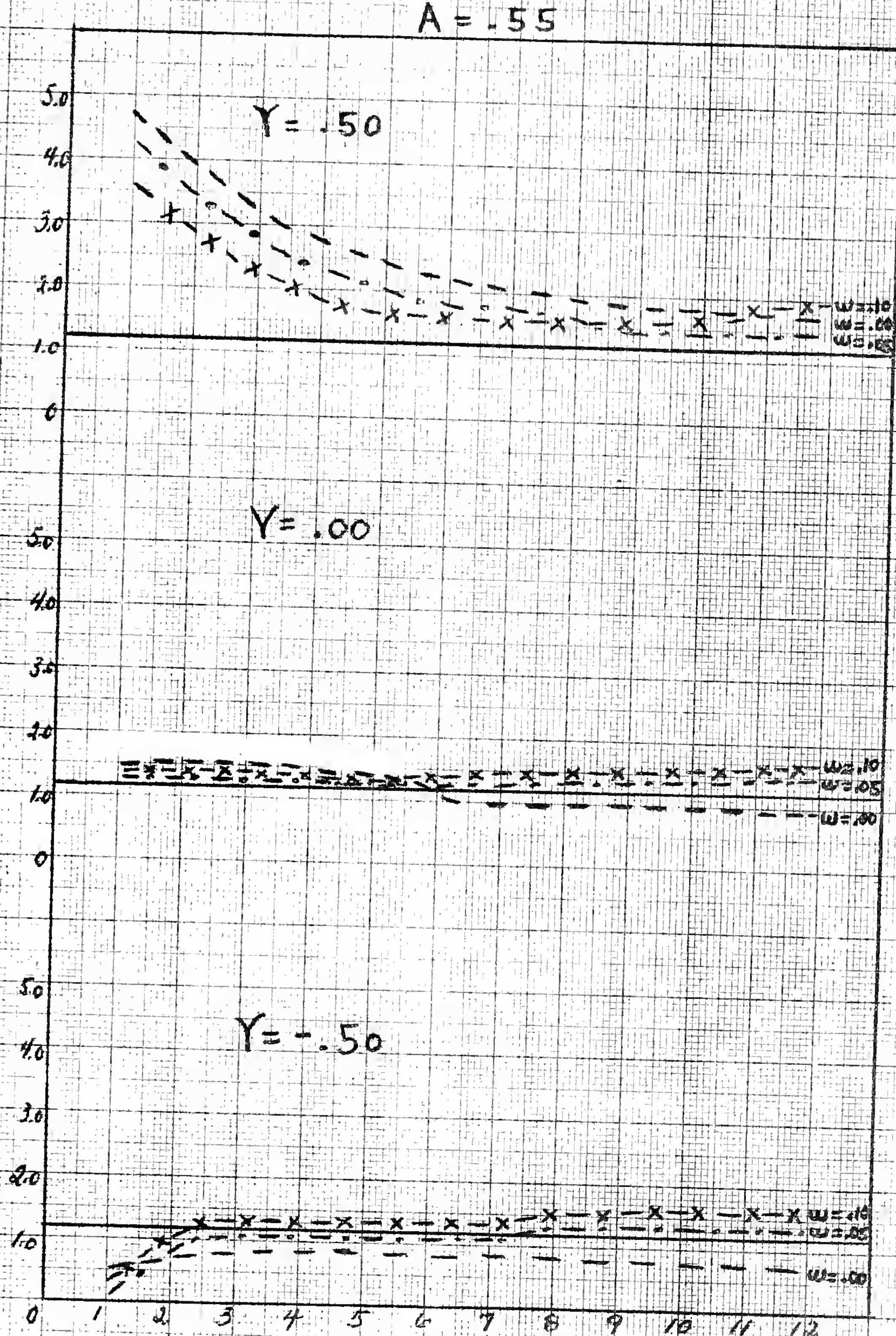

figure 1

the strength of the disturbance terms and the systematic terms in the expression (12) can be altered if M is different from .5.

In summary, there are three basic characteristics of the behavior of this model as it is aggregated over time. First, the average lag will increase or decrease if the serial correlation is respectively negative or positive. Second, when $\omega \neq 0$ the estimate of the lag will be greater, especially in highly aggregated models. And third, if α is close to one, the size of ω will be more important in determining the average lag in aggregated models.

We agreed in this study to compare the values of two parameters of the lag distribution, the average lag and the long run propensity. Fortunately, it is possible to show that if the approximation $\omega = 0$ is valid, that the long run propensity should be unaffected by time aggregation. Recalling that the long run propensity is $\hat{B}/1-\hat{A}$ and substituting for (5) we obtain

$$20) \qquad \text{Long Run Propensity} = \frac{(Y,Y)(Y,X) - (Y_{-n},X)(Y_{-n},Y)}{(Y,Y)(X,X)-(Y_{-n},X)^2-(X,X)(Y_{-n},Y)+(Y,X)(Y_{-n},X)}$$

Using the formulas in (L)) and observing that when $f_x(\theta)$ is a spike at the origin, $(Y_{-n},X) = (Y,X)$; (20) becomes simply

$$21) \qquad \text{Long Run Propensity} = \frac{(Y,X)}{(X,X)} = \frac{\Sigma\beta_i}{\Sigma\alpha_i}$$

where β_i and α_i are the coefficients of the rational polynomials in (1). Since (21) is true for all n and since this is also the Long Run Propensity for the true model (1), <u>time aggregation leaves the estimates of long run propensities consistent as long as the exogenous variables are primarily trend.</u> Thus the long run properties of a model should not be affected by time aggregation, only the short run properties are altered.

A simple generalization of the preceding set of results to Liu's [11] "inverted v" model is easily obtained. This lag distribution is obtained by replacing the exogenous variable by its own moving average of several periods with the consequence that the peak effect is felt not in the first period, but several periods later. The longer the moving average, the later the peak will occur. Such a model is very appealing from an econometric point of view, especially for models using short time periods.

In terms of the analysis in this paper, this change merely implies changing $\beta(L)$ and $B(L^n)$ to the appropriate length moving average assumed by the true model and the estimated model.

Letting $R_r(L) = \frac{1}{r^2}[1 + L + L^2 + \ldots + L^{r-1}]$ and $\tilde{X} = R_r(L^n)$, (3) becomes

$$22) \qquad Y = AY_{-n} + B\tilde{X} + W$$

and (5) is merely to be rewritten with $\sim$ over all X's. Similarly, in equations (10) the only change is the presence of $R_r(e^{in\theta})$ in the last two integrals, and (X,X) becomes

$$23) \qquad (\tilde{X},\tilde{X}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|R_r(e^{in\theta})\right|^2 \left|R_n(e^{i\theta})\right|^2 f_x(\theta)\ d\theta.$$

Because of the myriad possible combinations of true monthly moving average and aggregate moving average, only the approximation $\omega = 0$ will be presented. From the manner in which the $\beta(L)$ and $B(L^n)$ have been introduced into (5), it is clear that only the terms with $f_x(\theta)$ are altered, and that when $\theta = 0$, they remain identical to the strict Koyck-Nerlove model. There is a difference, however, in that the average lag implied by any estimate of A depends on the length of the moving average on X. If X is subjected to a moving average of r periods, then

the average lag is given by

$$24) \qquad \text{AV Lag} = n\left(\frac{A}{1-A} + \frac{r-1}{2}\right) .$$

Thus, in order to analyze such a model when aggregated over time, we merely subtract $\frac{n(r-1)}{2}$ from each estimate and compare the results in terms of the theory expounded above. That is,the theory should apply to the average lags computed without regard for lags in the exogenous variables.

It is perfectly clear that again, our approximation has taken much of the impact out of the exogenous variables but a careful study of the consequences for each pair of true and estimated moving average models would be very time consuming. For an analysis of the case where only the disaggregated model has a moving average see Engle [2] .

III. EMPIRICAL RESULTS

In order to ascertain whether the analytical results were reasonable and whether the approximations employed were helpful, seven equations from Liu's [11] monthly econometric model of the U.S. were reestimated in quarterly and annual forms using exactly the same data. The model which most closely approximated the form of the monthly model* was estimated in each case using ordinary least squares, two-stage least squares, and an estimator which consistently estimates a Koyck-Nerlove model in the presence of first or second order serial

*This means that the total time covered by moving averages have been kept the same and that one month lags in monthly exogenous variables have become simultaneous observations in the quarterly and annual case, while one month lags in the dependent variable have become one period lags in the aggregated version. It is possible that in some cases, a different form might give more reliable estimates but this paper will shed no light on this question.

correlation. The program, which will be called autoregressive least squares or ALS, was written by Martin [12]; it performs an autoregressive transformation and then estimates the model with non-linear constraints on the parameters. If indeed, the biases resulting from the introduction of serial correlation into the disturbance during aggregation are greater than simultaneous equation bias, then this estimator might prove better than the TSLS estimator in approximating the dynamic structure of the underlying model. The analytical theory however should best describe TSLS estimators since it assumes that the exogenous variables are independent of the disturbance.

The seven equations which are all taken directly from the Liu [11] model are Consumer Nondurables (CN), Consumer Services (CS), Business Construction (BC), Equipment (Q), Dividends (DIV), Corporate Profits (CP), and Corporate Profits after Depreciation (CPDC). These are the only equations which are estimated in the Koyck form; the complete listing of the regression results is in the appendix. Here we shall merely display the relevant statistics on the average lag and long run propensity, each computed with respect to the leading GNP variable.

In figure 3, the average lags for each equation are plotted against the level of aggregation. The values are the OLS monthly estimate and the TSLS quarterly and annual estimates each of which is adjusted for the inverted v formulation of the model. The horizontal line is the "true" value as given by Liu using a presumably consistent estimating technique. Finally, on each diagram, the ALS estimates are plotted with a star. The order of the diagram is that the highest serial correlation is plotted first.

The curves are in fact very much like those of figures 1 and 2. For positive serial correlation the average lag falls and for small and negative

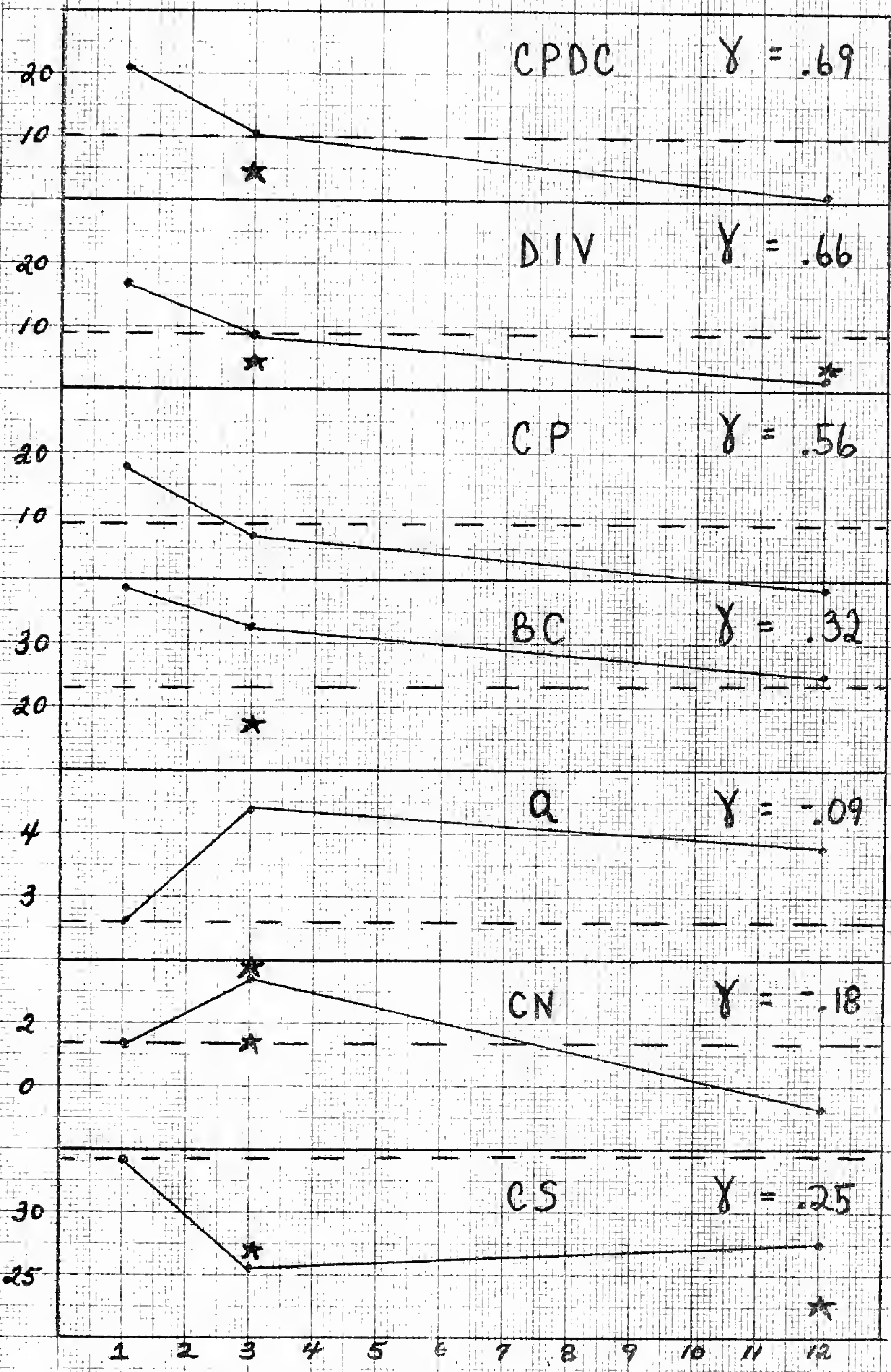

figure 3 AVERAGE LAG

serial correlation, it rises and then falls again. However, one equation, CS, does not have the proper shape. In theory one would expect the monthly estimate to be below the quarterly but this is not the case. In general the results are quite good and give strong corroboration to the approximations inherent in the theory. The ALS estimates are however not markedly better than the TSLS. As is shown in Engle and Liu [11], they are in fact much worse in terms of system response. The reasons for this may be twofold. First, simultaneous equation bias may dominate the biases due to serial correlation; but second and perhaps more important, the processes of the disturbances are complicated stochastic processes which cannot be represented by simple first and second order autoregressive models. Therefore, the assumption of this specification may in fact not improve the estimator at all for some stochastic disturbances and could have unforeseen deleterious effects.

The long run propensities are given in Table 1 and, although they are not the same with aggregation, the differences are not large. Almost nowhere are these different by more than a factor of 2 and this difference, in view of the standard errors is not significant.

Table I

Long Run Propensities

Equation	Ser. Corr.	Monthly	Quarterly	Annual
CN				
OLS	-.18	.093	.155	.099
TSLS			.097	.095
ALS			.201	.195
CS				
OLS	-.25	.430	.450	.480
TSLS			.442	.480
ALS			.445	.500

Equation	Ser. Corr.	Monthly	Quarterly	Annual
BC				
OLS	.32	.480	.486	.281
TSLS		.350	.465	.283
ALS			.312	
Q	-.09	.381	.712	.674
OLS			.735	.684
TSLS			.634	
ALS				
DIV				
OLS	.66	.242	.223	.210
TSLS		.213	.221	.209
ALS			.216	.208
CP				
OLS	.56	.269	.211	.167
TSLS		.194	.204	.167
ALS				
CPDC				
OLS	.69	.178	.136	.080
TSLS		.088	.139	.080
ALS			.094	

Thus in conclusion, the data in Liu's model provides reasonable corroboration of the theoretical predictions for the biases to be expected from aggregation over time of Koyck-Nerlove type distributed lag models. The econometrician can, from this result, infer the plausible size of the bias which might be inherent in an aggregated model. The analysis shows the sensitivity of even a well specified distributed lag model to changes in the unit time period and casts uncertainty on the dynamic properties of models with improper time units.

APPENDIX A

In practice, the simplest method for evaluating the complicated integrals in (10), is to use Cauchy's integral theorem and consider the integral of θ from $-\pi$ to π as a line integral around a unit circle in the complex plane. The value of any closed line integral is zero if the surface inside the path is everywhere analytic (most functions are analytic with as notable exceptions points of non-differentiability where the function becomes infinite). This observation corresponds to the fact that any round trip hike on a frictionless trail theoretically involves no net work. If however, there are singularities within the path, the integral will have a value which is easily computed. The value is given by Cauchy's theorem to be $2\pi i$ times the residue of the singularity where the residue is the coefficient of the singular factor. For example, $\frac{6X}{X-3}$ has a singularity at $X = 3$ with a residue of $6X$ or 18.

This theorem is easily used to evaluate the integrals of (10). For example, to evaluate

$$\text{A1)} \quad \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{i\theta} d\theta}{(1-\alpha e^{-i\theta})}$$

let $e^{+i\theta} = z$ giving

$$\text{A2)} \quad \frac{1}{2\pi i} \oint_{|z|=1} \frac{dz}{\left(1 - \frac{\alpha}{z}\right)} = \frac{1}{2\pi i} \oint_{|z|=1} \frac{z dz}{(z-\alpha)} \quad .$$

Now, if $|\alpha| < 1$, there is a singularity at $z = \alpha$ with value α which is within the unit circle and (A2) can easily be evaluated as

$$\text{A3)} \quad \frac{2\pi i\alpha}{2\pi i} = \alpha \text{ .}$$

Integrals of the form of (A1) arise whenever moments are computed between two variables which are lagged functions of a white noise process. Suppose $x = g(L)\varepsilon$ and $y = h(L)\varepsilon$, then

$$(x,y) = \Big(g(L)\varepsilon,\ h(L)\varepsilon\Big) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \sigma^2 \ g(e^{-i\theta})\ h(e^{+i\theta})\ d\theta,$$

and making the substitutions from above, but writing L instead of Z,

$$\text{A4)} \qquad (x,y) = \sigma^2 \sum_{\substack{\text{residue in}\\ \text{unit circle}}} g(L^{-1})h(L)L^{-1} \,.$$

However $(x,y) = (y,x)$ so

$$\text{A5)} \qquad (x,y) = \sigma^2 \sum_{\substack{\text{residue in}\\ \text{unit circle}}} g(L)h(L^{-1})L^{-1}$$

and therefore in general

$$\text{A6)} \qquad \sum_{\substack{\text{residue in}\\ \text{unit circle}}} f(L) = \sum_{\substack{\text{residue in}\\ \text{unit circle}}} f(L^{-1})\ L^{-2}$$

Equations (A4), (A5) and (A6) provide the tools necessary to evaluate the integrals which have continuous spectral density functions since these can be expressed in terms of a rational polymomial of the lag operator on a white noise process. It is apparent that if two singularities occur at the same point, the above formulas will not work. For the generalization to this case, see any advanced calculus book such as Kaplan [10].

As an example of the use of (A4) and (A6) let us evaluate the following integral from (10).

$$\text{A7)} \qquad \frac{1}{2\pi}\int_{-\pi}^{\pi} \left| R(e^{i\theta}) \right|^2 \frac{d\theta}{\left|1-\alpha e^{i\theta}\right|^e \left|1-\gamma e^{i\theta}\right|^2}$$

$$= \sum_{\text{residue}} \frac{R(L)R(L^{-1})}{(1-\gamma L)(1-\frac{\alpha}{L})(1-\alpha L)(1-\frac{\alpha}{L})} L^{-1}$$

$$= \sum_{\text{residue}} \frac{\frac{1}{n^2}[1+L+L^2+\ldots+L^{n-1}][1+\frac{1}{L}+\frac{1}{L^2}+\ldots+\frac{1}{L^{n-1}}]}{(1-\alpha L)(L-\alpha)(1-\gamma L)(L-\gamma)}$$

A8) $$= \sum_{\text{residue}} \frac{\frac{L}{n^2}[n+(n-1)(L+\frac{1}{L})+(n-2)(L^2+\frac{1}{L^2})+\ldots+(L^{n-1}+\frac{1}{L^{n-1}})}{(1-\alpha L)(L-\alpha)(1-\gamma L)(L-\gamma)}$$

This complicated function has singularities at $L = \alpha$, $L = \gamma$, $L = 0$ all of which are by assumption within the unit circle. Unfortunately, the singularity at $L = 0$ is a multiple root and therefore the recommended procedure is to evaluate part of this sum in this form and then apply (A6) in order to calculate the rest. Taking all the terms in the numerator with positive exponents first, the integral becomes

A9) $$\frac{1}{n^2}\left\{\frac{\alpha[n+(n-1)\alpha+\ldots+\alpha^{n-1}]}{(1-\alpha^2)(1-\gamma\alpha)(\alpha-\gamma)} + \frac{\gamma[n+(n-1)\gamma+\ldots+\gamma^{n-1}]}{(1-\alpha\gamma)(\gamma-\alpha)(1-\gamma^2)}\right\}$$

$$+ \sum_{\text{residue}} \frac{1}{n^2} \frac{L\left[(n-1)\frac{1}{L}+\ldots+\frac{1}{L^{n-1}}\right]}{(1-\alpha L)(L-\alpha)(1-\gamma L)(L-\gamma)}$$

Applying (A6) to the last term gives

$$\text{last term} = \sum_{\text{residue}} \frac{1}{n^2} \frac{1}{L} \frac{(n-1)L+\ldots+L^{n-1}}{(1-\frac{\alpha}{L})(\frac{1}{L}-\alpha)(1-\frac{\gamma}{L})(\frac{1}{L}-\gamma)} \frac{1}{L^2}$$

$$= \sum_{\text{residue}} \frac{1}{n^2} \frac{L[(n-1)L+\ldots+L^{n-1}]}{(L-\alpha)(1-\alpha L)(L-\alpha)(1-\gamma L)}$$

which is evaluated directly just as in (A9) to obtain as the full integral of (A7)

$$= \frac{\alpha Q(\alpha)}{(1-\alpha^2)(1-\gamma\alpha)(\alpha-\gamma)} + \frac{\gamma Q(\gamma)}{(1-\alpha\gamma)(\gamma-\alpha)(1-\gamma^2)}$$

using the definition of $Q(\alpha)$ following (12).

Appendix B†

Corporate Profits after Depreciation

Monthly OLS $\gamma = .560$

$$CPCD = \underset{(.017)}{.955}\ CPCD_{-1} + \underset{(.002)}{.008}\ GNP_{-1} + \underset{(.030)}{.161}\ PW_{-1} - 30.6$$

AV LAG†† (Adj)= 22.2 (21.2) LR Prop = .178 $R^2 = .975$

Monthly First Iteration

$$CPCD = \underset{(.034)}{.909}\ CPCD_{-1} + \underset{(.003)}{.008}\ GNP_{-1} + \underset{(.037)}{.041}\ PW_{-1} - 1.97$$

AV LAG (Adj) = 11.0 (10.0) LR Prop = .088 $R^2 = .883$

Quarterly TSLS

$$CPCD^{*} = \underset{(.067)}{.798}\ CPCD^{*}_{-3} + \underset{(.006)}{.028}\ GNP^{*} + \underset{(1.19)}{5.04}\ PW^{*} - 93.7$$

AV LAG (Adj) = 11.9 LR Prop = .139 $R^2 = .863$ (1.29)

Quarterly First Order $\gamma = .517$ (.322)

$$CPCD^{*} = \underset{(.279)}{.629}\ CPCD^{*}_{-3} + \underset{(.019)}{.035}\ GNP^{*} + \underset{(1.54)}{3.37}\ PW^{*} - 28.3$$

AV LAG (Adj) = 5.1 LR Prop = .094 $R^2 = .889$ (1.75)

Annual TSLS

$$CPCD^{**} = \underset{(.289)}{.079}\ CPCD^{**}_{-12} + \underset{(.021)}{.074}\ GNP^{**} + \underset{(4.82)}{7.40}\ PW^{**} - 122.7$$

AV LAG (Adj) = 1.0 LR Prop = .080 $R^2 = .543$ (2.08)

Dividends

Monthly OLS $\gamma = .658$

$$DIV = \underset{(.018)}{.942}\ DIV_{-1} + \underset{(.004)}{.014}\ CPCT^{*}_{-1} + .098$$

AV LAG (Adj) = 18.3 (16.3) LR Prop = .242 $R^2 = .995$

Monthly First Iteration

$$DIV = \underset{(.032)}{.906}\ DIV_{-1} + \underset{(.007)}{.020}\ CPCT^{*}_{-1} + .085$$

AV LAG (Adj) = 11.7 (9.7) LR Prop = .213 $R^2 = .981$

Quarterly TSLS

$$DIV^{*} = \underset{(.054)}{.747}\ DIV^{*}_{-3} + \underset{(.011)}{.056}\ CPCT^{*} + .482$$

AV LAG (Adj) = 8.6 LR Prop = .221 $R^2 = .981$ (1.58)

Quarterly First Order $\gamma = .267\ (.148)$

$$DIV^{*} = \underset{(.074)}{.689}\ DIV^{*}_{-3} + \underset{(.014)}{.067}\ CPCT^{*} + .50$$

AV LAG (Adj) = 6.6 LR Prop = .216 $R^2 = .982\ (2.03)$

Annual TSLS

$$DIV^{**} = \underset{(.173)}{.119}\ DIV^{**}_{-12} + \underset{(.033)}{.184}\ CPCT^{**} + 2.02$$

AV LAG (Adj) = 1.6 LR Prop = .209 $R^2 = .948\ (2.49)$

Annual First Order $\gamma = -.353\ (.322)$

$$DIV^{**} = \underset{(.188)}{.235}\ DIV^{**}_{-12} + \underset{(.034)}{.157}\ CPCT^{**} + 2.67$$

AV LAG (Adj) = 3.7 LR Prop = .208 $R^2 = .962\ (1.12)$

Corporate Profits

Monthly OLS $\gamma = .560$

$$CP = \underset{(.021)}{.948}\ CP_{-1} + \underset{(.004)}{.014}\ GNP_{-1} + \underset{(.032)}{.170}\ PW_{-1} - 33.3$$

AV LAG (Adj) = 19.2 (18.2) LR Prop = .269 $R^2 = .993$

Monthly First Iteration

$$CP = \underset{(.034)}{.897}\ CP_{-1} + \underset{(.006)}{.020}\ GNP_{-1} + \underset{(.041)}{.107}\ PW_{-1} - 9.26$$

AV LAG (Adj) = 9.7 (8.7) LR Prop = .194 $R^2 = .975$

Quarterly TSLS

$$CP^{*} = \underset{(.076)}{.720}\ CP^{*}_{-3} + \underset{(.013)}{.058}\ GNP^{*} + \underset{(1.18)}{5.18}\ PW^{*} - 100.8$$

AV LAG (Adj) = 7.7 LR Prop = .204 $R^2 = .966\ (1.23)$

Annual TSLS

$$CP^{**} = \underset{(.233)}{-.200}\ CP^{**}_{-12} + \underset{(.033)}{.200}\ GNP^{**} + \underset{(3.84)}{7.22}\ PW^{**} - 139.2$$

AV LAG (Adj) = -2.0 LR Prop = .167 $R^2 = .934\ (1.90)$

Business Construction

Monthly OLS $\gamma = .324$

$$BC = \underset{(.014)}{.975}\ BC_{-1} + \underset{(.004)}{.012}\ CPCT^{**}_{-1} \underset{(.042)}{-.101}\ R_{-1} + .28$$

AV LAG (Adj) = 45.5 (39.0) LR Prop = .480 $R^2 = .990$

Monthly First Iteration

$$BC = \underset{(.020)}{.958}\ BC_{-1} + \underset{(.006)}{.015}\ CPCT^{**}_{-1} \underset{(.059)}{-.097}\ R_{-1} + .27$$

AV LAG (Adj) = 29.3 (22.8) LR Prop = .350 $R^2 = .980$

Quarterly TSLS

$$BC^* = \underset{(.045)}{.916}\, BC^*_{-3} + \underset{(.013)}{.039}\, CPCT^{**} \; \underset{(.143)}{-.295}\, R^* + .77$$

AV LAG (Adj) = 37.2 (32.7) LR Prop = .465 R^2 - .963 (1.41)

Quarterly First Order γ = .365 (.161)

$$BC^* = \underset{(.086)}{.846}\, BC^*_{-3} + \underset{(.021)}{.048}\, CPCT^{**} \; \underset{(.209)}{-.231}\, R^* + .82$$

AV LAG (Adj) = 21.0 (18.0) LR Prop = .312 R^2 = .967 (2.02)

Annual TSLS

$$BC^{**} = \underset{(.209)}{.678}\, BC^{**}_{-12} + \underset{(.048)}{.091}\, CPCT^{**} \; \underset{(.710)}{-.396}\, R^{**} + 2.78$$

AV LAG (Adj) = 25.2 LR Prop = .283 R^2 = .828 (2.03)

Business Equipment

Monthly OLS γ = -.086

$$Q = \underset{(.066)}{.711}\, Q_{-1} + \underset{(.064)}{.110}\, CPCT^{**}_{-1} \; \underset{(.413)}{-1.58}\, R_{-1} + \underset{(.012)}{.026}\, t + 7.81$$

AV LAG (Adj) = 9.06 (2.56) LR Prop = .381 R^2 = .940

Quarterly TSLS

$$Q^* = \underset{(.095)}{.593}\, Q^*_{-3} + \underset{(.078)}{.299}\, CPCT^{**} \; \underset{(.564)}{-1.36}\, R^* \; \underset{(.036)}{-.005}\, t^* + 3.34$$

AV LAG (Adj) = 8.87 (4.37) LR Prop = .735 R^2 = .929 (1.42)

Quarterly First Order γ = .420 (.26)

$$Q^* = \underset{(.248)}{.382}\, Q^*_{-3} + \underset{(.154)}{.391}\, CPCT^{**} \; \underset{(1.08)}{-1.93}\, R^* + \underset{(.059)}{.015}\, t^* + 3.73$$

AV LAG (Adj) = 6.35 (1.85) LR Prop = .634 R^2 = .943 (1.73)

Annual TSLS

$$Q^{**} = \underset{(.231)}{.237}\, Q^{**}_{-12} + \underset{(.156)}{.522}\, CPCT^{**} \; \underset{(1.68)}{-.914}\, R^{**} + \underset{(.120)}{.062}\, t^{**} + 2.98$$

AV LAG (Adj) = 3.72 LR Prop = .684 R^2 = .869 (1.92)

Consumer Non-Durables

Monthly OLS γ = -.18

$$CN = \underset{(.06)}{.550}\, CN_{-1} + \underset{(.024)}{.042}\, Y^*_{-1} + \underset{(.008)}{.022}\, M + \underset{(.02)}{.282}\, t + 34.8$$

AV LAG (Adj) = 3.2 (1.2) LR Prop = .093 R^2 = .992

Quarterly TSLS

$$CN^* = \underset{(.106)}{.545}\, CN^*_{-3} + \underset{(.038)}{.044}\, Y^* + \underset{(.012)}{.022}\, M^* + \underset{(.095)}{.283}\, t^* + 34.6$$

AV LAG (Adj) = 3.6 LR Prop = .097 R^2 = .994 (2.12)

Quarterly Second Order $\gamma 1 = .63\ (.186)$ $\gamma 2 = -.51\ (.139)$

$$CN^{*} = \underset{(.237)}{.320}\ CN^{*}_{-3} + \underset{(.053)}{.137}\ Y^{*} + \underset{(.018)}{.009}\ M^{*} + \underset{(.165)}{.310}\ t^{*} + 20.9$$

AV LAG (Adj) = 1.4 LR Prop = .201 $R^2 = .995\ (2.04)$

Annual TSLS

$$CN^{**} = \underset{(.275)}{-.082}\ CN^{**}_{-12} + \underset{(.086)}{.102}\ Y^{**} + \underset{(.029)}{.035}\ M^{**} + \underset{(.329)}{.703}\ t^{**} + 86.5$$

AV LAG (Adj) = -.9 LR Prop = .095 $R^2 = .994\ (1.95)$

Consumer Services

Monthly OLS $\gamma = -.248$

$$CS = \underset{(.017)}{.972}\ CS_{-1} + \underset{(.007)}{.012}\ Y^{*}_{-1} + \underset{(.002)}{.004}\ M\ -1.19$$

AV LAG (Adj) = 36.6 (34.6) LR Prop = .430 $R^2 = .999$

Quarterly TSLS

$$CS^{*} = \underset{(.041)}{.896}\ CS^{*}_{-3} + \underset{(.018)}{.046}\ Y^{*} + \underset{(.006)}{.009}\ M^{*}\ -4.17$$

AV LAG (Adj) = 25.8 LR Prop = .442 $R^2 = .999(2.43)$

Quarterly First Order $\gamma = -.217\ (.125)$

$$CS^{*} = \underset{(.032)}{.899}\ CS^{*}_{-3} + \underset{(.014)}{.045}\ Y^{*} + \underset{(.005)}{.009}\ M^{*}\ -4.9$$

AV LAG (Adj) = 26.6 LR Prop = .445 $R^2 = .999\ (2.08)$

Annual TSLS

$$CS^{**} = \underset{(.067)}{.696}\ CS^{**}_{-12} + \underset{(.030)}{.146}\ Y^{**} + \underset{(.011)}{.016}\ M^{**}\ -13.3$$

AV LAG (Adj) = 27.5 LR Prop = .480 $R^2 = .999\ (1.60)$

Annual First Order $\gamma = .293\ (.274)$

$$CS^{**} = \underset{(.082)}{.653}\ CS^{**}_{-12} + \underset{(.047)}{.174}\ Y^{**} + \underset{(.019)}{.008}\ M^{**}\ -11.3$$

AV LAG (Adj) = 22.6 LR Prop = .500 $R^2 = .999\ (1.94)$

---

† The variables are defined exactly as in Liu [11] where one star is a three month moving average and two stars is a twelve month moving average. Subscripts refer to months and the average lags are all computed in months. The variables which are not named by their equation are as follows:

GNP: Gross national product.
PW: PW = (P/W) / (GNP) P: GNP deflator, W: Salary and wage payments.
CPCT: Gross corporate profits and inventory valuation adjustments less corporate profit tax liability.
R: Interest rate.
t: Time trend (initial quarter = 1).
Y: Personal disposable income.
M: Personal liquid assets in billions of 1958 dollars.

††Formulas for computing average lag

$$\text{Av Lag}\left[\frac{b}{1-aL}\right] = \frac{a}{1-a}$$

$$\text{Av Lag}\left[\frac{b\,\frac{1}{r}\,(1 + L + L^2 + \ldots + L^{r-1})}{1 - aL}\right] = \frac{a}{1-a} + \frac{r-1}{2}$$

$$\text{Av Lag}\left[w(L^n)\right] = n\ \text{Av Lag}\left[w(L)\right]$$

$$\text{Av Lag}\left[L^m\, w(L)\right] = \text{Av Lag}\left[w(L)\right] + m$$

Bibliography

[1] Doob, J.L. Stochastic Processes. New York: John Wiley, 1953.

[2] Engle, Robert F. "Biases from Time Aggregation of Distributed Lag Models," Ph.D. Dissertation, Cornell University, 1969.

[3] Engle, Robert F. and T.C. Liu. "Effects of Aggregation over Time on Dynamic Characteristics of an Econometric Model," NBER, Conference on Research in Income and Wealth, November 1969.

[4] Goldberger, Arthur. Econometric Theory. New York: John Wiley, 1964.

[5] Granger, C.W.J. "The Typical Spectral Shape of an Economic Variable," Econometrica, Vol. 34, No. 1, p. 150, 1966.

[6] Griliches, Zvi. "Distributed Lags: A Survey," Econometrica, Vol. 35, 1967.

[7] Hannan, E.J. and R.D. Terrell, "Testing for Serial Correlation After Least Squares Regression," Econometrica, Vol. 36, No. 1, January 1968, pp. 133-150.

[8] Howrey, E.P. "Stochastic Properties of the Klein-Goldberger Model," forthcoming in Econometrica.

[9] Ironmonger, D.S. "A Note on the Estimation of Long Run Elasticities," Journal of Farm Economics, XLI, 1959, pp. 626-32.

[10] Kaplan, Wilfred. Advanced Calculus. Reading: Addison-Wesley, 1959.

[11] Liu, T.C. "A Monthly Recursive Econometric Model of the United States: A Test of Feasibility," Review of Economics and Statistics, Vol. 51, February 1969.

[12] Martin, James. "Computer Algorithms for Estimating the Parameters of Selected Classes of Non-Linear Single Equation Models," P - 585, Oklahoma University, 1968.

[13] Mundlak, Yair. "Aggregation over Time in Distributed Lag Models," International Economic Review, May 1961, Vol. 2, No. 2, p. 154.

[14] Nerlove, Marc. "On the Estimation of Long Run Elasticities, A Reply," Journal of Farm Economics, XLI, 1959, pp. 632-40.

[15] Nerlove, Marc. "Spectral Analysis of Seasonal Adjustment Procedures," Econometrica, Vol. 32, No. 3, p. 241, 1964.

[16] Sims, Christopher. "Discrete Approximations to Continuous Time Distributed Lags in Econometrics," forthcoming in Econometrica.

[17] Telser, Lester. "Discrete Samples and Moving Sums in Stationary Stochastic Processes," Journal of the American Statistical Association, Vol. 62, No. 318, p. 484, 1967.